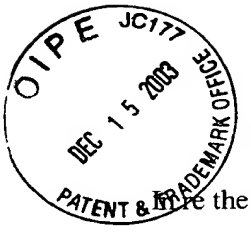
AF/2654

Xerox Docket No. D/A0466

PATENT APPLICATION

2700



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE HONORABLE BOARD OF PATENT APPEALS AND INTERFERENCES

In re the Application of

Salah AIT-MOKHTAR et al.

On Appeal from Group: 2654

Application No.: 09/738,319

Examiner: V. Harper

Filed: December 18, 2000

Docket No.: 108169

For: METHOD AND APPARATUS FOR GENERATING NORMALIZED REPRESENTATIONS OF STRINGS

APPEAL BRIEF TRANSMITTAL

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Attached hereto are three (3) copies of our Brief on Appeal in the above-identified application.

The Commissioner is hereby authorized to charge Deposit Account No. 24-0037 in the amount of Three Hundred Thirty Dollars ($330.00) in payment of the Brief fee under 37 C.F.R. 1.17(f). In the event of any underpayment or overpayment, please debit or credit our Deposit Account No. 24-0037 as needed in order to effect proper filing of this Brief.

For the convenience of the Finance Division, two additional copies of this transmittal letter are attached.

Respectfully submitted,

Robert J. Webster

James A. Oliff
Registration No. 27,075

Robert J. Webster
Registration No. 46,472

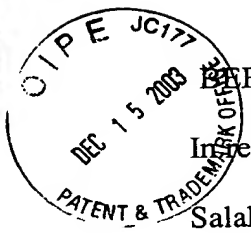
JAO:RJW/sxb
Date: December 15, 2003

OLIFF & BERRIDGE, PLC
P.O. Box 19928
Alexandria, Virginia 22320
Telephone: (703) 836-6400

DEPOSIT ACCOUNT USE AUTHORIZATION
Please grant any extension necessary for entry;
Charge any fee due to our Deposit Account No. 24-0037

PATENT APPLICATION

#12



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE HONORABLE BOARD OF PATENT APPEALS AND INTERFERENCES

In re the Application of:

Salah AIT-MOKHTAR et al.

Application No.: 09/738,319

Filed: December 18, 2000 Docket No.: 108169

For: METHOD AND APPARATUS FOR GENERATING NORMALIZED REPRESENTATIONS OF STRINGS

BRIEF ON APPEAL

Appeal from Group 2654

OLIFF & BERRIDGE, PLC
P.O. Box 19928
Alexandria, Virginia 22320
Telephone: (703) 836-6400
Attorneys for Appellants

TABLE OF CONTENTS

I. INTRODUCTION

This is an Appeal from a final rejection mailed July 15, 2003, finally rejecting claims 1-8 and 10-20. No claims are allowed.

A. Real Party in Interest

The real party in interest for this Appeal in the present application is Xerox Corporation by way of an Assignment recorded at Reel 11402 Frame 0183.

Statement of Related Appeals and Interferences

There are presently no appeals or interferences, known to Appellants, Appellants' representative or the Assignee, which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

B. Status of Claims

Claims 1-8 and 10-20 are pending. Claims 1-8 and 10-20 stand finally rejected and are on appeal.

C. Status of Amendments

1. Claim 9 was canceled and claims 19 and 20 were added by a Preliminary Amendment filed on February 26, 2001. The Preliminary Amendment has been entered.

2. Claim 10 was amended in an Amendment filed on June 23, 2003. This Amendment has been entered.

II. SUMMARY OF THE INVENTION AND APPLIED REFERENCES

A. The Invention

The invention generally relates to a method and apparatus for generating normalized representations of strings, e.g. sentences, and in particular to a method and apparatus for providing translation information for translating a string from a first language to a second language.

The invention has various objects, including providing a method and apparatus: for generating normalized representations of strings; for increasing the proportion of relevant text units retrieved from a database; for retrieving information from a database with a high precision; and for generating normalized representations which may be stored in a database wherein the normalized representations increase the retrieval performance of information retrieval systems. The invention also provides a translation memory for translation systems having an increased retrieval performance of translation information, as well as a method for providing translation information with improved retrieval performance from previously translated text units.

According to the invention, a method for normalizing input strings includes receiving the input strings for linguistic analysis, which may include morphological, syntactic and semantic analysis, to generate a first representation of each of the input strings. Each of the first representations includes linguistic information that may relate, for example, strings such as sounds and words to their meaning and vice versa. Each of the first representations is skeletised to generate a corresponding second representation for each of the input strings. The skeletising replaces the linguistic information with abstract variables in each of the second representations. The second representation are then stored as normalized representations of the input strings.

The invention also provides a retrieval method for retrieving strings from a database. The database comprises normalized representations of a plurality of strings. The retrieval method includes receiving a query, generating a normalized representation of the query, matching the normalized representation of the query to the normalized representations stored in the database and retrieving strings from the database that are identified by the matching step.

The invention also provides translation information for translating a string of a first language to a second language. The translation information is based on a plurality of normalized representations of strings in the first language and the second language. The method for providing translation information includes receiving an input string in the first language, retrieving a similar string in the first language from the plurality of normalized representations and outputting the translation information based on a string in the second language which corresponds to the retrieved string in the first language.

Additional processing steps are also applied, either by providing at least one operation out of a plurality of operating functions, extracting linguistic information, or by operating functions that generalize extracted linguistic information. The invention has applications in indexing, information retrieval, translation memory, authoring memory and machine edited human translation.

The selection of skeletons for sentences depends on the application for which the normalization procedure is intended. Skeletons relevant for translation memories may not be relevant for retrieval or for monolingual authoring. For retrieval applications, it is important that a retained skeletisation function efficiently partitions the set of sentences, so that the search for similar sentences involves few steps. Thus, skeletisation functions which discriminate best-input sentences and which stop skeletising sentences when the associated skeletisation function applies only to a few sentences are to be selected. Similar considerations apply to other kinds of applications. A selection of operating functions, in particular skeletisation functions, and their order may be carried out, for example, by machine learning.

In Fig. 1 of the application, the information retrieval system 101 includes a "search line" which comprises a query formatting unit 104 for formatting a query and generating a normalized representation for the query. In the formatting stage, the query is transformed by

the unit 104 into a normalized representation, using the same methods 103 as described for text units. The results of retrieving operations performed by a matching unit 106 are output as retrieved strings.

Fig. 4 of the application illustrates in more detail the processing of strings, i.e. text units or sentences, for obtaining a database according to the invention. The manner of knowledge representation, namely architecture of the database, is an essential aspect of the invention. The particular knowledge representation and method of generating such knowledge representation is not restricted to a document retrieval application, but may be a basis for knowledge extraction, a translation information apparatus, etc.

Turning to the process described in Fig. 4 of the application, in step 401, linguistic information is extracted from an input sentence as described above. Linguistic information may include morphological, syntactical, and semantic information and ways to disambiguate between them. Based on the extracted information, a normalized representation is generated.

Such a normalizing procedure may be carried out a plurality of times with respect to different aspects. The performed sequence of operations will result in a corresponding plurality of normalized representations having a different degree of generalization.

Further, as indicated in step 402, a skeletisation operation is applied to the normalized representation(s). The skeletisation operation replaces particular elements of a string or sentence by a variable in order to allow for better matches during an information retrieval process. Such variables will match to corresponding variables of other sentences regardless of the particular originally used text elements in text units or sentences to be matched.

Each, or only selected ones, of the normalized representations resulting from the linguistic extracting step and the skeletisation step may be stored together with the particular sentence, as for example strings 102.

B. The Claims

Independent claim 1 recites a method of normalizing input strings comprising (a) receiving the input strings; (b) linguistically analyzing the input strings to generate a first representation of each of the input strings, each of the first representations including linguistic information; (c) skeletising each of the first representations to generate a corresponding second representation for each of the input strings, said skeletising step replacing the linguistic information with abstract variables in each of the second representations; and (d) storing the second representation as normalized representations of the input strings.

Dependent claims 2-8 and 19-20 further define the step of linguistically analyzing the input strings recited in claim 1.

Dependent claim 10 further defines the abstract variables recited in claim 1.

Dependent claim 11 further defines the normalized representations recited in claim 1.

Dependent claim 12 recites steps in addition to the steps recited in claim 1.

Dependent claims 13 and 14 further define steps (a)-(d) recited in claim 1.

Independent claim 15 recites an apparatus for normalizing input strings, comprising a text processing unit for (a) receiving the input strings, (b) linguistically analyzing the input strings to generate a first representation of each of the input strings; each of the first representations including linguistic information, and (c) skeletising each of the first representations to generate a corresponding second representation for each of the input strings; said skeletising replacing the linguistic information with abstract variables in each of the second representations; and (2) memory for storing the second representation as normalized representations of the input strings.

Dependent claim 16 further recites a query formatting unit.

Dependent claim 17 further recites a memory and a matching unit.

Dependent claim 18 further recites a translation memory.

C. **U.S. Patent No. 6,006,221 to Liddy**

Liddy discloses a multilingual document retrieval system using semantic vector matching. Each document in a database is subjected to a set of processing steps to generate a language-independent conceptual representation of the subject content of the document. The query is subjected to a set of processing steps to generate a language-independent conceptual representation of the subject content of the query. The documents and queries can be subjected to additional analysis to provide additional term-based representations. Documents are matched to queries based on the conceptual-level content of the documents. The query's representation is compared to each documents' representation to generate a measure of relevance of the document to the query. See the Abstract of Liddy.

Liddy's processing modules, shown in Fig. 2, are discussed in terms of (1) preprocessor 110, LI 120, POS tagger 130, and PNC 140 that perform initial processing for tagging and identification; (2) MCGRE 150, MCGD 160, MCG-MHCM 170, MHCD 180, and MCVG 190 that generate conceptual-level representations of the documents and queries; (3) PTI 210 and PQP 220 that generate term-based representations of the documents and queries; (4) MCVM 200 and QDM and score combiner 230 that correlate the document and query information to provide an evaluation of the documents; and (5) recall predictor 240 and GUI 250 that present the results to the user.

The meaning of these acronyms is found in column 7, lines 25-47, of Liddy.

Liddy performs (1) initial processing and tagging (discourse level tagging and SGML-like tagging) (col. 8, lines 40-53); (2) language identification (col. 8, line 54 to col. 9, line 6); (3) part of speech tagging (col. 9, line 7 to col. 9, line 18); (4) proper noun identification and categorization (col. 9, line 19 to col. 10, line 49); (5) conceptual level representation (col. 10, line 50 to col. 16, line 14); (6) generation of term-based representations (col. 16, line 15 to

col. 18, line 29); (7) matching of documents with queries (col. 18, line 30 to col. 19, line 55); (8) and presentation of results (col. 19, line 56 to col. 22, line 53).

D. Collins, "Discriminative Reranking for natural language Parsing," Proc. 17th International Conference on machine Learning," July 2000.

Collins discloses alternative methods that re-rank the output of an existing machine-language statistical parser. The base parser produces a set of candidate parses for each input sentence, with associated probabilities that define an initial ranking of those parses. A second model then attempts to improve on the initial ranking of those parses.

III. THE ISSUES ON APPEAL

1. Whether, under 35 USC §102(e), claims 1-8, 10-18 and 20 are anticipated by Liddy.

2. Whether, under 35 U.S.C. §103(a), claim 19 would have been obvious over Liddy in view of Collins.

IV. GROUPING THE CLAIMS ON APPEAL

All claims do not stand or fall together.

Group I: Claims 1-8, 9-18 and 20 all include the skeletising feature of step (c) recited in claim 1, and are patentable for all the reasons expressed below.

Group II: Claim 19 includes the features of claims 1 and 2 and further recites the feature of performing machine learning for selecting particular operation functions out of a plurality of operating functions and for determining the processing order. Claim 19 is patentable for all the reasons stated below.

There are two independent claims, i.e., claims 1 and 15, and 17 dependent claims, identified below:

Claims 2, 10, 11 and 13 depend directly from independent claim 1.

Claims 3-8, 14, 19 and 20 depend indirectly from independent claim 1.

Claim 16 depends directly from independent claim 15.

Claims 17 and 18 depend indirectly from independent claim 15.

V. THE LAW

A. 35 USC §102 (Anticipation)

A prior art reference anticipates the subject of a claim when the reference discloses every feature of the claimed invention, either explicitly or inherently (see, Atlas Powder Co. v. IRECO Inc., 190 F.3d 1342, 1346, 51 USPQ2d 1943, 1945-46 (Fed. Cir. 1999); In re Schreiber, 128 F.3d 1473, 1477, 44 USPQ2d 1429, 1431(Fed. Cir. 1997); In re Paulsen, 30 F.3d 1475, 1478,1479, 31 USPQ2d 1671, 1675 (Fed. Cir. 1994), In re Spada, 911 F.2d 705, 708, 15 USPQ2d 1655, 1657 (Fed. Cir. 1990), Hazani v. Int'l Trade Comm'n, 126 F.3d 1473, 1477, 44 USPQ2d 1358, 1361 (Fed. Cir. 1997); RCA Corp. v. Applied Digital Data Systems, Inc., 730 F.2d 1440, 1444, 221 USPQ 385, 388 (Fed. Cir. 1984); In re Schreiber, 128 F.3d 1473, 1477, 44 USPQ2d 1429, 1431 (Fed. Cir. 1997); and Hazani v. Int'l Trade Comm'n, 126 F.3d 1473. 1477. 44 USPQ2d 1358, 1361 (Fed Cir. 1997)). While, of course, it is possible that it is inherent in the operation of the prior art device that a particular element operates as theorized by the examiner, inherence may not be established by probabilities or possibilities. In re Oelrich, 666 F.2d 578, 581, 212 USPQ 323, 326 (CCPA 1981) and In re Rijckaert, 9 F.3d 1531, 1534, 28 USPQ2d 1955, 1957 (Fed. Cir. 1993).

It is well settled that the burden of establishing a prima facie case of anticipation resides with the Patent and Trademark Office (PTO). See, In re Piasecki, 745 F.2d 1468, 223 USPQ 785, 788 (Fed. Cir. 1984). When relying upon the theory of inherency, the Examiner must provide a basis in fact and/or technical reasoning to reasonably support the determination that the allegedly inherent characteristic necessarily flows from the teachings of the applied prior art. See, Ex parte Levy, 17 USPQ2d 1461, 1464 (Bd. Pat. App. & Int. 1990). Inherency is a question of fact, In re Schreiber, supra, In re Fracalossi, 681 F.2d 792, 794, 215 USPQ 569, 571 (CCPA 1982). To establish inherency, the extrinsic evidence

must make clear that the missing descriptive matter is necessarily present in the thing described in the reference, and that it would be so recognized by persons of ordinary skill; inherency, however, may not be established by probabilities or possibilities. The mere fact that a certain thing may result from a given set of circumstances is not sufficient. See Continental Can Co. v. Monsanto Co., 948 F.2d 1264, 1268-69, 20 USPQ2d 1746, 1749 (Fed. Cir. 1991).

An inventor may choose to be his or her own lexicographer by defining, with reasonable clarity, deliberateness and precision, the specific terms used to describe his invention. In re Paulsen, 30 F.3d 1475, 1480, 31 USPQ2d 1671, 1674 (Fed. Cir. 1994). In this regard, words which are defined in the specification must be given the same meaning when used in a claim. McGill, Inc. v. John Zink Co., 736 F.2d 666, 674, 221 USPQ 944, 949 (Fed. Cir. 1984), cert. denied, 469 U.S. 1037.

Moreover, all positively recited features of a claim must be addressed in an Office Action. In re Angstadt et al., 190 USPQ 214 (CCPA 1976); In re Alul et al., 175 USPQ 700 (CCPA 1972); and General Electric Company v. U.S., 198 USPQ 65 (US ClCt 1978).

B. 35 U.S.C. §103(a) (Obviousness)

The Supreme Court in Graham v. John Deere, 383 U.S. 1 at 18, 148 USPQ 459 at 467 (1966), set forth the basic test for patentability under 35 U.S.C. §103:

Under §103, the scope and content of the prior art are to be determined; differences between the prior art and the claims at issue are to be ascertained; and the level of ordinary skill in the pertinent art resolved. Against this background, the obviousness or nonobviousness of the subject matter is determined. Such secondary considerations as commercial success, long felt but unresolved need, failure of others, etc., might be utilized to give light to the circumstances surrounding the origin of the subject matter to be patented.

Moreover, in In re Ehrreich and Avery, 200 USPQ 504, 509-510 (CCPA 1979), the Court of Customs and Patent Appeals further clarified the basic test set forth in Graham v. John Deere:

We must not here consider a reference in a vacuum, but against the background of the other references of record which may disprove theories and speculations in the reference or reveal previously undiscovered or unappreciated problems. The question in a §103 case is what the references would collectively suggest to one of ordinary skill in the art. In re Simon, 461 F.2d 1387, 174 USPQ 114 (CCPA 1972). It is only by proceeding in this manner that we may fairly determine the scope and content of the prior art according to the mandate of Graham v. Deere Company, 383 US 1, 17, 148 USPQ 459, 467 (1966). (Emphasis in original.)

Thus, the mere fact that parts of prior art disclosures can be combined does not make the combination obvious unless the prior art also contains something to suggest the desirability of the combination. In re Imperato, 486 F.2d 585 (CCPA 1973).

To imbue one of ordinary skill in the art with knowledge of the invention, when no prior art reference or references of record convey or suggest that knowledge, is to fall victim to the insidious effect of hindsight syndrome wherein that which only the inventor taught is used against its teacher. W.L. Gore & Assoc. v. Garlock, Inc., 721 F.2d 1540, 1533, 220 USPQ 303, 312-13 (Fed. Cir. 1983).

Further, analyzing the claimed invention as a whole in view of the prior art as a whole, one indicium of nonobviousness is a "teaching away" from the claimed invention by the prior art at the time the invention was made. See U.S. v. Adams, 148 USPQ 479 (1966). Essentially, teaching away from a claimed invention is a per se demonstration of lack of prima facie obviousness.

Where the prior art provides "only general guidance and is not specific as to the particular form of the invention or how to achieve it, [such a suggestion] may make an approach 'obvious to try,' but it does not make the invention obvious." Ex parte Obukowicz, 27 USPQ2d, 1063, 1065 (U.S. Patent and Trademark Office Board of Appeals and Interferences, 1992); and In re O'Farrell, 7 USPQ2d 1673, 1681 (Fed. Cir. 1988).

Further, in In re Wright, 848 F.2d 1216, 6 USPQ2d 1959 (Fed. Cir. 1988), the Federal Circuit stated:

Factors including unexpected results, new features, solution of a different problem, novel properties are all considerations in the determination of obviousness. . . .

These secondary considerations (objective evidence of non-obviousness), as outlined in Graham v. John Deere and further characterized in In re Wright, must be evaluated before reaching an ultimate decision under 35 U.S.C. §103.

The test for obviousness is what the combined teachings would have suggested to one of ordinary skill in the art. See, In re Young, 927 F.2d 588, 591, 18 USPQ2d 1989, 1091 (Fed. Cir. 1991) and In re Keller, 642 F.2d 413, 425, 208 USPQ 871, 881 (CCPA 1981). More specifically, as stated by the court in Keller, 642 F.2d at 425, 208 USPQ at 881, the test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary references; nor is it that the claimed invention must be expressly suggested in one or all of the references. Rather, the test is what the combined teachings of the references would have suggested to those of ordinary skill in the art. Moreover, the artisan is not compelled to blindly adopt every single aspect of the teachings of any one reference without the exercise of independent judgment See Lear Siegler, Inc. v. Aeroquip Corp., 733 F.2d 881, 889, 221 USPQ 1025, 1032 (Fed. Cir. 1984).

With regard to motivation to combine the references used in the rejection of appellants' claims, while there must be some teaching, reason, suggestion or motivation to

combine existing elements to produce the claimed device, it is not necessary that the cited references or prior art specifically suggest making the combination. See, B.F. Goodrich Co. v. Aircraft Braking Systems Corp., 72 F.3d 1577, 1583, 37 USPQ2d 1314, 1319 (Fed. Cir 1996) and In re Nilssen, 851 F.2d 1401, 1403, 7 USPQ2d 1500, 1502 (Fed. Cir. 1988). Rather, the test for obviousness is what the combined teachings of the references would have suggested to one of ordinary skill in the art.

The Office Action must provide proper motivation to combine the teaching of different references. The first requirement of proper motivation is that a showing of a suggestion, teaching, or motivation to combine the prior art references is an "essential evidentiary component of an obviousness holding." C.R. Bard, Inc. v. M3 Sys. Inc., 157 F.3d 1340, 1352, 48 USPQ2d 1225, 1232 (Fed. Cir. 1998). This evidence may flow from the prior art references themselves, the knowledge of one of ordinary skill in the art, or, in some cases, from the nature of the problem to be solved. See Pro-Mold & Tool Co. v. Great Lakes Plastics, Inc., 75 F.3d 1568, 1573, 37 USPQ2d 1626, 1630 (Fed. Cir. 1996). However, the suggestion more often comes from the teachings of the pertinent references. See In re Rouffet, 149 F.3d 1350, 1359, 47 USPQ2d 1453, 1459 (Fed. Cir. 1998). This showing must be clear and particular, and broad conclusory statements about the teaching of multiple references, standing alone, are not "evidence." See Dembiczak, 175 F.3d at 1000, 50 USPQ2d at 1617.

The Office Action must also demonstrate that modifying one reference in view of another reference is even feasible. Moreover, the case law requires that for motivation to be proper, showing that something is feasible is not enough. Just because something is feasible does not mean that it is desirable or that one of ordinary skill in the art would be motivated to do what is feasible. See Winner International Royalty Corp. v. Wang, 53 USPQ2d 1580

(Fed. Cir. 2000), which points out that motivation to combine references requires a showing not just of feasibility, but also of desirability.

In Tec Air Inc. v. Denso Manufacturing Michigan Inc., 52 USPQ2d 1294 (Fed. Cir. 1999), the Court of Appeals for the Federal Circuit stated that there is no suggestion to combine relevant teachings from different references if a reference teaches away from its combination with another source. The court also stated that a reference may be said to teach away when a person of ordinary skill in the art, upon reading the reference, would be discouraged from following the path set out in the reference, or would be led in a direction divergent from the path that was taken by the applicant.

Additionally, in In re Braat, 16 USPQ2d 1812 (Fed. Cir. 1990) (unpublished), the Court of Appeals for the Federal Circuit reversed a decision by the PTO Board of Appeals and Interferences, stating that the reference upon which the obviousness of claim 1 was based taught away from the claimed invention, and that "[O]ne important indicium of non-obviousness is "teaching away" from the claimed invention by the prior art," citing In re Dow Chemical Co., 5 USPQ2d 1529, 1532 (Fed. Cir., 1988).

Moreover, a factual inquiry whether to modify a reference must be based on objective evidence of record, not merely conclusionary statements of the Examiner. See, In re Lee, 277 F.3d 1338, 1343, 61 USPQ2d 1430, 1433 (Fed. Cir. 2002).

VI. ARGUMENT

A. The Rejections

Claims 1-8, 10-18 and 20 are rejected under 35 USC §102(e) as anticipated by U.S. Patent No. 6,202,064 to Liddy.

This rejection alleges that Liddy discloses features which "correspond" to the claimed features.

Claim 19 is rejected under 35 USC §103(a) as unpatentable over Liddy in view of the Collins non-patent literature reference. The Office Action admits that Liddy fails to disclose "performing machine learning for selecting particular operating functions out of said plurality of operating functions and for determining the processing order." The Office Action alleges that "[I]n the same field of endeavor, Collins teaches the machine learning technique of discriminative reranking for natural language parsing where reranking techniques can be applied to problems in natural language processing to improve the resulting representations. . . ." The Office Action then concludes that it would be obvious "to modify Liddy by specifically providing reranking techniques, as taught by Collins" for the purpose of improving the resulting representation."

B. Liddy does not anticipate Claims 1-8, 10-18 and 20

Liddy generates a language-independent conceptual representation of the subject content of a document and can subject the document to additional analysis to provide term-based representations, such as extraction of information-rich terms and phrases, such as proper nouns. See the Summary of the Invention in column 2 of Liddy.

The final rejection was repeated verbatim from its form in the previous (May 15, 2003) Office Action, despite the fact that in the "Response to Arguments" portion of the final rejection, a completely different part of Liddy (col. 15, line 63 through col. 16, line 10) was relied upon to allegedly disclose the feature of "skeletising each of the first representations to generate a corresponding second representation for each of the input strings; said skeletising step replacing the linguistic information with abstract variables in each of the second representations" than was relied upon in the final rejection as stated. The stated final rejection relies on column 6, lines 15-20, Figs. 1 and 2, column 6, line 63 through column 7, line 5 of Liddy to allegedly disclose this feature.

Applicants respectfully requested clarification of the actual basis of the rejection insofar as the "skeletising" feature of claims 1-8, 10-18 and 20 is concerned, and received a statement on pages 2 and 3 of the Advisory Action that the reference to column 15, line 63, to column 16, line 10 of Liddy, was merely to clarify the rejection. Applicants address both the original basis of the final rejection as well as the alleged clarification of that basis, below.

To the extent that the Office Action asserts that this feature is anticipated by the disclosure of column 6, lines 15-20, Figs. 1 and 2, and column 6, line 63 - column 7, line 5 as "generating both conceptual and term-based alternative representations of the documents and queries with relevant information extracted from the documents and indexed," Applicants respectfully disagree.

While Liddy linguistically analyzes input strings to generate first representations (of those strings) which include linguistic information, Liddy does not perform skeletising of each of the first representations to generate a corresponding second representation for each of the input strings, the skeletising step replacing the linguistic information with abstract variables in each of the second representations, as recited in claim 1.

In Liddy, the "analogous processing" mentioned in column 6, lines 15-20 is said to be "to determine the requirements for document matching." The alternative representations of the documents and queries are said to be both conceptual and term based. See lines 19 and 20 of column 6 of Liddy. However, Liddy never performs further processing that obtains second representations obtained by replacing the linguistic information obtained by the initial processing with abstract variables. Nor does Liddy provide a system to do so.

Fig. 2 of Liddy shows a preferred method of operation, for example, in which a part of speech tagger 130 is disclosed as outputting a parts of speech tagged document and a proper noun identifier and a categorizer 140 is disclosed as identifying and tagging proper nouns. Modules 150-190 generate monolingual vector codes of the subject contents of both

documents and queries. Module 150 tags each word with the codes of all multilingual concept groups to which various senses of the word/phrase in the document belong. Module 160 outputs a fully tagged text stream with a single multilingual concept group for each word in the input text. Module 170 outputs a tagged, native language text stream with unique, monolingual (English), hierarchical categories assigned to each substantive word. Module 180 outputs a text stream with disambiguated monolingual categories assigned to each substantive word. Module 190 produces a fixed-dimension vector representation of the concept-level context of the text.

None of this processing in Liddy replaces earlier performed linguistic analysis. Liddy performs additional processing of a query or of a document, but does not replace the earlier processing results. Not only does Liddy fail to disclose replacing linguistic information with other information, but Liddy also fails to disclose replacing linguistic information with abstract variables.

With respect to the additional reasons presented in the "Response to Arguments" section of the final Office Action, i.e., with respect to Fig. 5 and column 15, line 63 to column 16, line 10 of Liddy, Applicants respectfully submit that the invention recited in claim 1 is not disclosed in those portions of Liddy.

As discussed above, Liddy does not replace linguistic information with an "abstract variable" as recited in claim 1. The Microsoft Press Computer Dictionary, 1991 edition, defines "abstract data type" as "a data type that is defined in terms of the information it can contain and the operations that can be performed with it. An abstract data type is more generalized than one constrained by the properties of the object it contains . . ."

Applicants respectfully submit that what is disclosed in the paragraph bridging columns 15 and 16 of Liddy is constrained by the properties of the object it contains, i.e., the codes to which individual words are mapped clearly represent certain properties (e.g.,

meanings) of those individual words. Therefore, the codes of Liddy are not abstract data of any type, let alone "abstract variables", as recited in claim 1. There is no indication in Liddy that the codes are defined in terms of the information they can contain and in terms of the operations they can perform. In other words, the disambiguated concept codes of Liddy are not "abstract variables" as recited in claim 1.

Moreover, the "Response to Arguments" portion of the final rejection does not address what "abstract variables" are. Instead, the final rejection simply states that a word is disambiguated and then represented, i.e., replaced with disambiguated concept codes. Thus, the final Office Action does not present any evidence to demonstrate that "disambiguated concept codes" correspond to "abstract variables" as recited in the claims.

In this regard, the "Response to Arguments" portion of the Advisory Action, on page 4, alleges that Liddy generates both conceptual and term-based alternative representations of the documents and queries, citing column 6, lines 17-19, including the processing performed in the MCGD step, citing column 15, line 64 through column 16, line 10). The Advisory Action continues by alleging that, in addition, the use of "concept groups" and "concept categories", citing Fig. 2, clarified in Fig. 5, can be argued to be "abstract variables" because they are data types defined in terms of the information that they can contain.

Applicants point out that the Advisory Action is continuing to do what it has done throughout the prosecution of this application, i.e., switch the portion of Liddy being relied on once the portion previously relied on has been demonstrated by Applicants not to support the rejection.

The Advisory Action apparently has withdrawn the position taken in the final rejection that appeared to indicate that "disambiguated concept codes" correspond to the recited "abstract variables."

The Advisory Action now appears to focus on "conceptual and term-based alternative representations" but fails to explicitly state what these representations have to do with the claimed invention, leaving Applicants wondering what portions of Liddy the rejection is actually based on, and what features of the claims that the "conceptual and term-based alternative representations" are supposed to anticipate.

Moreover, the Advisory Action now asserts that "concept groups" and "concept categories" can be argued to be "abstract variables." Unfortunately, no reasoning or argument is presented to support this conclusion. Applicants are thus left to guess at the reasoning behind this unsupported conclusion.

Applicants respectfully submit that the continued switching from one ground of rejection to another demonstrates the lack of substantive and procedural due process (required by the Administrative Procedures Act) accorded Applicants during the prosecution of this application. This is reason enough to reverse the rejection. See in this regard, In re Zurko, 119 S.Ct. 1816, 50 USPQ2d 1930 (1999), and In re Gartside, 53 USPQ2d 1769 (Fed. Cir. 2000).

Turning to the merits of this conclusionary argument, Applicants respectfully submit that nothing in Figs. 2 and 5 of Liddy discloses skeletising each of the first representations to generate a corresponding second representation for each of the input strings, said skeletising step replacing the linguistic information with abstract variables in each of the second representations, as recited in claim 1.

Liddy merely classifies the string, e.g., places it in a generic category, such as "Finance," but does not replace the linguistic information with an abstract variable, as recited in claim 1.

Therefore, Liddy does not anticipate claim 1 and the final Office Action fails to make out a prima facie case of anticipation of claim 1 by Liddy. Claims 2-8, 10-18 and 20 depend

from claim 1 and, thus, contain the features thereof. Accordingly, Liddy does not anticipate claims 2-8, 10-18 and 20. Therefore, the rejection of claims 1-8, 10-18 and 20 under 35 USC §102(e) over Liddy is improper and should be reversed.

C. **Liddy and Collins do not render claim 19 obvious**

The final rejection rejects claim 19 under 35 USC §103(a) over Liddy in view of Collins. This rejection is improper and should be reversed.

Liddy does not anticipate claim 1 for the reasons stated above. Collins does not, and was not applied to, provide the features of claim 1 missing from Liddy as discussed above. Accordingly, the rejection is improper because, even if properly combined, the asserted combination of references does not disclose, teach or suggest every feature recited in claim 1.

Furthermore, Collins is cited to render obvious "performing machine learning for selecting particular operating functions out of said plurality of operating functions and for determining the processing order." Collins is alleged to teach the machine language technique of discriminative re-ranking for natural language parsing and to provide motivation in terms of an improved representation that results [in] improved recall and precision and decreased error. The alleged motivation to combine these references is for the purpose of "an improved representation that results [in] improved recall and precision and decreased error."

The "Response to Arguments" portion of the final Office Action presents a different rationale than was presented in the previous Office Action to allegedly provide motivation to modify Liddy to achieve the invention recited in claim 19. The "Response to Arguments" section of the final Office Action switches from one statement of what Collins discloses to a different statement of what Collins discloses as a basis for the alleged motivation to combine these references in the manner suggested.

The rejection itself states "Collins teaches the machine learning technique of discriminative re-ranking for natural language parsing where re-ranking techniques can be

applied to problems in natural language processing to improve the resulting representations (§1, "Introduction")." - emphasis added. However the "Response to Arguments" section states something different, by stating that "Collins teaches the machine learning technique of discriminative re-ranking for natural language parsing and gives a motivation in terms of an improved representation that results [in] improved recall and precision, and decreased error (§1, last paragraph)." - emphasis added.

Applicants respectfully submit that the Office Action fails to make out a prima facie case of obviousness for the reasons stated above concerning the shortcomings of Liddy, and because the Office Action fails to provide proper motivation to combine these references as alleged.

A showing of a suggestion, teaching, or motivation to combine the prior art references is an "essential evidentiary component of an obviousness holding." C.R. Bard, Inc. v. M3 Sys. Inc., 157 F.3d 1340, 1352, 48 USPQ2d 1225, 1232 (Fed. Cir. 1998). This evidence may flow from the prior art references themselves, the knowledge of one of ordinary skill in the art, or, in some cases, from the nature of the problem to be solved. See Pro-Mold & Tool Co. v. Great Lakes Plastics, Inc., 75 F.3d 1568, 1573, 37 USPQ2d 1626, 1630 (Fed. Cir. 1996). However, the suggestion more often comes from the teachings of the pertinent references. See In re Rouffet, 149 F.3d 1350, 1359, 47 USPQ2d 1453, 1459(Fed. Cir. 1998). This showing must be clear and particular, and broad conclusory statements about the teaching of multiple references, standing alone, are not "evidence." See In re Dembiczak, 175 F.3d 994 at 1000, 50 USPQ2d 1614 at 1617.

The assertion that one of ordinary skill in the art would have been properly motivated to combine these two references (i.e., found it desirable) for the purpose of achieving "an improved recall and precision" is not a clear and particular evidentiary teaching, but is only a broad conclusory statement. All information search and retrieval systems are evaluated in

terms of precision, which is the percentage of relevant documents retrieved to the total number of documents retrieved, and recall, which is the percentage of the relevant documents retrieved to the total number of relevant documents in the database searched. Thus, the asserted motivation is generic at least.

Moreover, Collins is concerned with natural language parsing which attempts to determine the sequence structure of sentences, whereas Liddy is concerned with categorizing individual words, and does not parse sentences. The Office Action never explains how one is allegedly motivated to modify Liddy's word-by-word translation system with Collins' sentence parsing system, or what in Liddy is re-ranked and how it is re-ranked, or what results from such an allegedly desirable re-ranking. Applicants respectfully submit that the details of how Liddy is allegedly modified by Collins are left up to speculation or further invention.

The Advisory Action addresses these arguments by asserting that:

> Liddy teaches a method for multilingual document retrieval which includes a set of processing steps (sequence of functions) to generate a language-independent conceptual representation of the subject content of a document (text) (title, abstract, Fig. 2). Collins teaches the use of machine learning techniques to improve the representation of natural language processing (where natural language processing is a sequence of operations to generate a representation of text) (§1, "Introduction"). Thus the functions are similar.

Unfortunately, the Advisory Action fails to explain (1) in what way "the functions are similar;" or (2) why, based on that unexplained similarity, one of ordinary skill in the art would have the desire to modify Liddy "by specifically providing reranking techniques;" or

(3) what reranking techniques have to do with "performing machine learning for selecting particular operation functions out of said plurality of operating functions and for determining the processing order."

Additionally, the Advisory Action, on page 5, attempts to clarify the statement in the final Office Action that the motivation to modify Liddy by specifically providing re-ranking techniques, as taught by Collins, for the purpose of improving the resulting representation, by asserting that "it is well-known in the art that an improved representation can reduce memory requirements."

Applicants' response to this clarification is that the alleged well-known result is speculative at best, and is nothing more than a broad conclusory statements about the teaching of multiple references, standing alone, and, as such, is not "evidence." See In re Dembiczak, 175 F.3d 994 at 1000, 50 USPQ2d 1614 at 1617 (Fed. Cir. 1999).

Finally, even if these two references were somehow properly combined, they would not render obvious the method of claim 1 because they are directed to different functions (e.g., word-by-word translation versus sentence parsing) and have different objects (e.g., document retrieval versus natural language parsing). In fact, the words "parse" and "parsing" are not found anywhere in Liddy.

Accordingly, the final rejection fails to make out a prima facie case of obviousness of the invention recited in claim 19. Therefore, the rejection of claim 19 under 35 USC §103(a) as unpatentable over Liddy in view of Collins is improper and should be reversed.

D. Rebuttal of Arguments in Advisory Action

The Examiner's rebuttal arguments vary from Office Action to Office Action, as pointed out above. In addition to the arguments set forth above, Applicants address the rebuttal arguments not specifically addressed above.

With respect to the issue of whether the ground of rejection in the final rejection was chanted from the non-final rejection, which is discussed on pages 2 and 3 of the Advisory Action, the Examiner's arguments are directed to petitionable matter. Applicants have not chosen to pursue this issue by petition the Commissioner under 37 CFR §1.181, so the issue is moot. Instead, Applicants have chosen to address the merits of the final rejection, as discussed above.

The issue of whether Liddy fails to disclose replacing linguistic information with abstract variables is addressed above. However, on page 3 of the Advisory Action, the Examiner states the belief "that Liddy does disclose the replacement of linguistic information with other information." Applicants note that claim 1 does not simply recite replacing linguistic information with other information. Instead, claim 1 recites, among other features, "replacing the linguistic information with abstract variables in each of the second representations." This positively recited feature is not even addressed on page 3 of the Advisory Action.

VII. CONCLUSION

Liddy does not disclose the features recited in claims 1-8, 10-18 and 20.

Neither Liddy nor Collins discloses the features recited in claim 19. One of ordinary skill in the art would not have been motivated to modify Liddy in view of Collins as alleged. Even if Liddy and Collins were properly combined, they would not result in, or render obvious, the invention recited in claim 19. The Office Action has not made out a prima facie case of anticipation of the subject matter recited in claims 1-8, 10-18 and 20, or a prima facie case of obviousness of the invention recited in claim 19 for the reasons stated above. Therefore, the final rejections of claims 1-8 and 10-20 should be reversed.

The Honorable Board is requested to reverse the rejections set forth in the final Office Action and to pass this application to issuance.

Respectfully submitted,

James A. Oliff
Registration No. 27,075

Robert J. Webster
Registration No. 46,472

JAO:RJW/sxb

Date: December 15, 2003

DEPOSIT ACCOUNT USE AUTHORIZATION
Please grant any extension necessary for entry;
Charge any fee due to our Deposit Account No. 15-0461

OLIFF & BERRIDGE, PLC
P.O. Box 19928
Alexandria, Virginia 22320
Telephone: (703) 836-6400

Enclosure:
Appendix

APPENDIX A

CLAIMS:

1. (Original) A method for normalizing input strings, the method comprising the steps of:

(a) receiving the input strings;

(b) linguistically analyzing the input strings to generate a first representation of each of the input strings; each of the first representations including linguistic information;

(c) skeletising each of the first representations to generate a corresponding second representation for each of the input strings; said skeletising step replacing the linguistic information with abstract variables in each of the second representations; and

(d) storing the second representation as normalized representations of the input strings.

2. (Original) The method of claim 1, wherein said step of linguistically analyzing comprises performing a plurality of operating functions.

3. (Original) The method of claim 2, wherein said plurality of operating functions comprise performing one of morphological analysis, syntactic analysis, and semantic analysis.

4. (Original) The method of claim 3, wherein said step of linguistically analyzing comprises normalizing words according to their base forms.

5. (Original) The method of claim 3, wherein said analysis further comprises the step of extracting a syntactic category for individual words.

6. (Original) The method of claim 3, wherein said analysis further comprises the step of extracting syntactic information representing string structure.

7. (Original) The method of claim 3, wherein said analysis further comprises the step of extracting dependency relations between sub-structures of a string.

8. (Original) The method of claim 3, wherein said analysis further comprises providing semantic links for individual words.

9. (Cancelled)

10. (Currently Amended) The method of claim 1, wherein the abstract variables are tags indicating the replaced linguistic information.

11. (Original) The method of claim 1, wherein the normalized representations are stored in a database.

12. (Original) The method of claim 11, further comprising: receiving a query; generating a normalized representation of said query by performing steps (b) and (c); matching the normalized representation of said query to the normalized representations stored in the database; and retrieving from said database strings identified by said matching step.

13. (Original) The method of claim 1, wherein said steps (a) - (d) are performed to generate a translation memory comprising a plurality of normalized representations of strings in a first language and a second language.

14. (Original) The method of claim 13, further comprising the steps of:

receiving an input string in the first language;

retrieving a similar string in said first language from said plurality of normalized representations, and

outputting said translation information based on a string in said second language which corresponds to said retrieved string in said first language.

15. (Original) An apparatus for normalizing input strings, the apparatus comprising:

a text processing unit for: receiving the input strings,

linguistically analyzing the input strings to generate a first representation of each of the input strings; each of the first representations including linguistic information, and

skeletising each of the first representations to generate a corresponding second representation for each of the input strings; said skeletising replacing the linguistic information with abstract variables in each of the second representations; and

memory for storing the second representation as normalized representations of the input strings.

16. (Original) The apparatus of claim 15, further comprising a query formatting unit for:

receiving queries;

linguistically analyzing the queries to generate a first representation of each of the queries; each of the first representations including linguistic information; and

skeletising each of the first representations to generate a corresponding second representation for each of the queries; said skeletising replacing the linguistic information with abstract variables in each of the second representations.

17. (Original) The apparatus of claim 16, further comprising:

memory for storing the second representation of the queries as normalized representations of the queries;

a matching unit for matching the normalized representations of the input strings with the normalized representation of the queries.

18. (Original) The apparatus of claim 16, further comprising a translation memory for storing translations of the input strings.

19. (Previously Added) The method of claim 2, further comprising the step of performing machine learning for selecting particular operation functions out of said plurality of operating functions and for determining the processing order.

20. (Previously Added) The method of claim 2, wherein said storing further comprises storing the operating functions performed on the normalized representations.